Parallel tutorial session #7 : MOLECULAR MECHANICS

08/09/2006

Wonmuk Hwang

Molecular mechanics fundamentals

- macroscopic mechanics (classical mechanics) : position, velocity, acceleration
  - describe the motion : kinematics
  - force = cause of motion : dynamics
  - few-body problem : deterministic from Newton's equation

  molecular mechanics (statistical mechanics)
  - many body problem → continuum approximation sometimes
  - stochastic, random, chaotic ≈ thermal noise
  - collective behavior can be known, details can be ignored to some extent

- the ensemble concept : one "macro-state" (structure, energy function, ...)
  ≡ many equally accessible "micro-states"
  (backbone chain orientations, bond lengths, ...)

  ensemble = a set of microstates that correspond to a given macrostate in equilibrium.

  molecular force : water and proteins are a closed system of very many microstates.
  - the macrostate of water is negligibly affected by the protein
  - water ≡ heat reservoir

  why does a protein bend (kinesin)? deterministic ≠ random motion

- entropy S

  Let Ω be the number of microstates in an ensemble ;

  $$\left\{ \begin{array}{l} \Omega = \exp(S/k_B) \\ S = k_B \ln \Omega \\ k_B = 1.38 \times 10^{-23}\ \mathrm{J/K} \\ \text{Boltzmann's constant} \end{array} \right.$$

  why is S defined as a logarithm? S is additive

  e.g. dice $\Omega_d = 6$, coin $\Omega_c = 2$

  both dice and coin $\Omega_{tot} = \Omega_d \times \Omega_c$

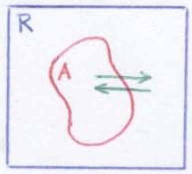
  $S_{tot} = S_d + S_c$

  some additive physical variables : E, N, V, S extensive quantities

  others are independent of system size : T, P, $\rho = N/V$ intensive quantities

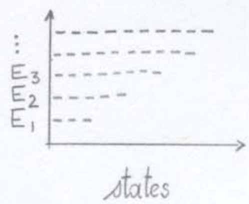
  S ↗ (entropy) when E ↗ (energy)

Boltzmann factor



R heat reservoir
A system of interest
$\rightleftarrows$ exchange of energy

Let $p_i = p(E_i)$ the probability that A is in a microstate of energy E_i



$$p_i \propto \Omega_R (E_{tot} - E_i) \quad \text{with } E_{tot} = E_R + E_i = \text{constant} \quad \text{(closed system)}$$

$$\Omega_R (E_{tot} - E_i) = \exp\left[S_R (E_{tot} - E_i) \div k_B \right]$$

$$S_R = S_R (E_{tot}) - E_i \left.\frac{\partial S_R}{\partial E}\right|_{E_{tot}} \qquad \text{because } E_i \ll E_{tot}$$

$$\frac{1}{T} \equiv \left.\frac{\partial S_R}{\partial E}\right|_{E_{tot}} \quad \text{defines temperature, hence} \quad S_R = S_R^{(tot)} - \frac{E_i}{T}$$

$$\boxed{p_i \propto \exp\left(\frac{-E_i}{k_B T}\right) = \exp(-\beta E_i)} \qquad \text{Boltzmann's factor}$$

after normalization $\quad p_i = \frac{1}{Z} \exp(-\beta E_i)$ with $\beta \equiv \frac{1}{k_B T}$

$$\boxed{Z = \sum_i \exp(-\beta E_i)} \qquad \text{partition function}$$

free energy G

start with partition function $\quad Z = \sum_i e^{-\beta E_i} = e^{-\beta F} \quad$ or $\quad \ln Z = -\beta F$

Helmholtz free energy $\quad F = - k_B T \ln Z \quad$ (in joules J)

we can now write $\quad p_i = \frac{1}{Z} \exp(-\beta E_i) \quad$ (in log:) $\ln p_i = -\beta E_i + \beta F$

and consider a particular state whose energy is equal to the average of the system A:

$$\bar{E} = \langle E \rangle$$

$$E_i = \bar{E} \Rightarrow \ln p(\bar{E}) = -\beta \bar{E} + \beta F = \overline{-\beta E + \beta F} = \overline{\ln p_i}$$

$$\langle \ln p_i \rangle = -\beta \bar{E} + \beta F = \sum_i p_i \ln p_i$$

Let Ω_i be the number of microstates with energy E_i ; $\quad p_i = \frac{1}{\Omega_i} = \exp\left(-\frac{S_i}{k_B}\right)$

from $\ln p_i = -\frac{S_i}{k_B}$, we get $\sum_i p_i \ln p_i = -\frac{1}{k_B} \sum_i p_i S_i$

and $\bar{S} = \sum_i p_i S_i \quad$ measurable entropy of the system

$$\frac{1}{k_B T}(-\bar{E} + F) = \frac{-\bar{S}}{k_B} \qquad \text{or} \qquad F = E - TS$$

$$\boxed{F = E - TS}$$

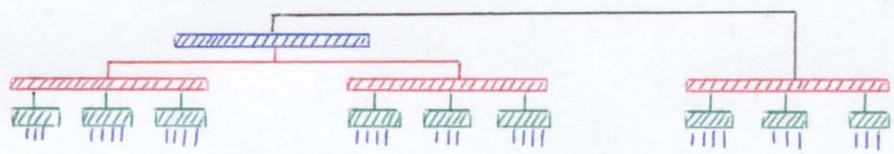
F: free energy of the system
E: internal energy of the system
S: entropy of the system
T: temperature of the reservoir

- Meaning:

- $F = E - TS$ "free" energy = maximum work the system can do
  let W be the work done by the system $\Delta W \leqslant -\Delta F$
  with $\Delta F \leqslant 0$

- $e^{-\beta F} = \sum_i e^{-\beta E_i}$ is a "representative Boltzmann's factor"
  it lumps up all Boltzmann's factors of the system.

protein conformation
electronic states :
nuclei quarks :



$e^{-\beta F'}$

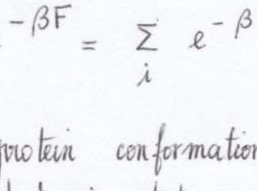
$e^{-\beta F}$

free energy enough to describe the whole system
at many time, length, energy scales

Gibbs free energy $G = H - TS$

- Equipartition theorem

in a typical equilibrium situation : fluctuations near (free) energy at equilibrium



$$E(x) = E(x_0) + \underbrace{(x - x_0)\, E'\big|_{x=x_0}}_{0 \text{ at min}} + \frac{(x-x_0)^2}{2} E''\big|_{x-x_0}$$

$$E(x) = E(x_0) + a\,(x - x_0)^2$$

$$p(x) = \frac{\exp\left(-\beta a (x-x_0)^2\right)}{Z} \qquad \text{with} \quad Z = \int_{-\infty}^{+\infty} dx \; e^{-\beta a (x-x_0)^2}$$

average energy :

$$\langle E \rangle = \int dx \; p(x)\; a\,(x-x_0)^2 = \frac{1}{Z} \int dx \; a\,(x-x_0)^2 \exp\left[-\beta a (x-x_0)^2\right]$$

$$= \frac{k_B T}{2} \qquad \text{using the Gaussian integral} \quad \int_{-\infty}^{+\infty} e^{-ay^2} dy = \sqrt{\frac{\pi}{a}}$$

energy per degree of freedom $= \frac{RT}{2}$ with $R = \mathcal{N}_A \, k_B$

if N degrees of freedom $\boxed{\langle E \rangle = \frac{N}{2} k_B T}$

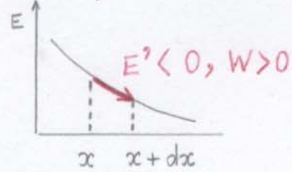
ex:  $N = 3 + 2$

$\bar{E} = \frac{5}{2} k_B T$

- Generalized force

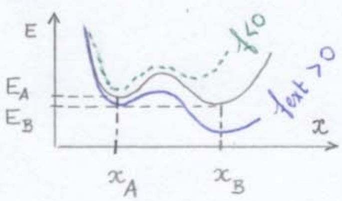
$E(x) = \dots$, x is called the control parameter

if the system does work W resulting in $x \to x + dx$



$$\left.\begin{aligned} dE &= dx \cdot \frac{\partial E}{\partial x} = -W \\ &= -f \, dx \end{aligned}\right\} \qquad \boxed{f = -\frac{\partial E}{\partial x}}$$

- So, in a two-state equilibrium



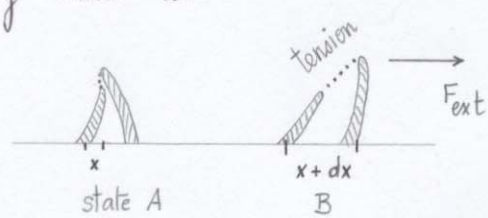
work done ON the system by an external force $f_{ext} \Delta x$

" " BY " " $= - f_{ext} \Delta x$

$E \longrightarrow E - f_{ext} \Delta x$

equilibrium constant $K_{eq} = \frac{[B]}{[A]} = e^{-\beta(\Delta E - f_{ext} \Delta x)}$ force tilts energy profile

in auditory hair cells



under force, " open state B " is made much more likely.

SUMMARY

statistical mechanics link micro to macro.

Molecular mechanics : ab initio foundation

Ju Li

see Powerpoint handout by instructor.

- quantum mechanics do matter to study hydrogen bonds
  haemoglobin / porphyrin interactions
  need to be understood to parameterize / approximate systems.

- why are electrons treated as quantum objects?

electrons $\{ \underline{x}_i \}$ $i = 1 .. n$ of mass m_e

ions $\{ \underline{\bar{X}}_I \}$ $I = 1 .. N$

$M_I \approx A_I \, m_p$, $m_p \approx 2000 \, m_e$

$A_I = 12$ and $Z_I = 6$

Molecular mechanics - 5.

$$V_{iI} = \frac{e^2}{|x_i - X_I|}$$ electrostatic interaction between electron & ion

de Broglie wavelength $\lambda = \frac{h}{|\bar{p}|}$, how large is it with respect to electrons?

kinetic energy $\frac{1}{2} m_e v_e^2 = \frac{p_e^2}{2 m_e} \sim 13.6 \text{ eV} \sim \frac{p_I^2}{2 m_I}$

for C - C bond $p_e \sim \sqrt{2 m_e * 13.6 \text{ eV}}$

$\lambda_e = 3.3$ Å

if $I \equiv C$ $\lambda \sim 0.02$ Å

$I \equiv H$ $\lambda \sim 0.08$ Å

" smearing factor "

position fluctuations, quantum, (not thermal !)

electrons's λ permeates through several bonds

↳ structural units

- Born- Oppenheimer approximation : ions are immobile
  electrons fluctuate

$$V_{BO}\left(\{\bar{X}_I\}_{I=1..N}\right) \equiv \min_{\Psi(\{x_i\})} E_{tot}\left(\{x_i\}, \{X_I\}\right)$$

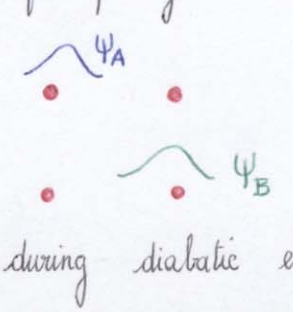
energy function $\Psi(\{x_i\}) = \Psi_G\left(\{x_i\} \mid \{X_I\}\right)$ electrons in ground state

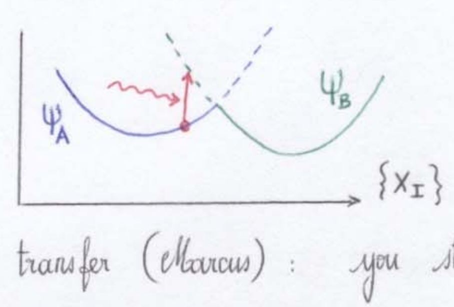
addendum : Newton's law $M_I \ddot{X}_I = - \frac{\partial V_{BO}(\{X_I\})}{\partial X_I} = \underline{F}_I$

steepness of energy landscape = force

when does the BO approximation break down ?

- if optically excited





light ⇝ can make electrons leave the least energetic state

- during diabatic electron transfer (Marcus) : you stay on Ψ_{cst} most of the time

when does the addendum break down ?

- at low temperature, or for low - weight ions (hydrogen) ≡ quantum too
  quantum fluctuation ≈ thermal fluctuation at low T

Eigenfunction problem $\hat{H}\,\Psi_G(x_1, \dots, x_n) = V_{BO}\,\Psi_G(x_1, x_2, \dots, x_n)$

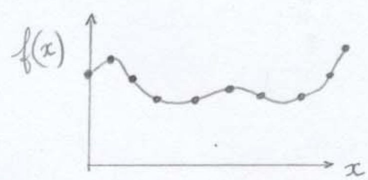
Hamilton operator $$\hat{H} = \sum_{i=1}^{n} \frac{-\hbar^2 \nabla_i^2}{2 m_e} + \sum_{i \neq j} \frac{e^2}{2|x_i - x_j|} - \sum_{i,I} \frac{Z_I e^2}{|x_i - X_I|} + \sum_{I \neq J} \frac{Z_I Z_J e^2}{2|X_I - X_J|}$$

$\left\{ \begin{array}{l} \text{solve with only parameters being Plank's constant } \hbar, m_e, e, Z_I \\ \text{ab initio approach} \end{array} \right.$

but information explosion: how to store $\Psi_G(x_1, y_1, z_1, \dots, x_n, y_n, z_n)$

if 1 variable 10 spline points of 8 Bytes each = 80 bytes

if 60 variables 10^{60} points
(20 electrons) or 10^{52} CDs of 100 MB each
!!!



fortunately, symmetry relations, and redundancy in spline functions

$\left. \begin{array}{l} \text{quantum chemistry (Pople)} \\ \text{density-functional theory (Kohn)} \end{array} \right\}$ address this issue of huge numbers.
error in bond energy calculations ↓

↳ advances in theoretical methodology
computational techniques
computing technology.

- Hartree-Fock theory

$$\Psi \approx S = \frac{1}{\sqrt{n!}} \begin{vmatrix} \tilde{\Psi}_1(x_1) & \cdots & \tilde{\Psi}_n(x_1) \\ \vdots & & \\ \tilde{\Psi}_1(x_n) & \cdots & \tilde{\Psi}_n(x_n) \end{vmatrix}$$

approximation of wavefunction by Slater determinant
from $3n$-dimensional to n 3-dimensional problem
add electron spin indication: ↑ or ↓ only takes 1 bit extra.

decompose the total energy into sum of pairs

$$\sum_{i,j} \Delta E(\Psi_i, \Psi_j) \quad \text{with} \quad \Delta E = E_{Hartree} - E_{exchange}$$

$$E_H(\tilde{\Psi}_i, \tilde{\Psi}_j) \equiv e^2 \iint dx\, dx' \frac{[\tilde{\Psi}_i^*(x)\, \tilde{\Psi}_i(x)][\tilde{\Psi}_j^*(x')\, \tilde{\Psi}_j(x')]}{|x - x'|} \quad \text{classical}$$

$E_{exchange}$ is quantum !

the $E_{exchange}$ stabilizes occupation of same spin wavefunction
(whereas spin doesn't matter in $E_{Hartree}$)
in fact, use several Slater determinants and minimize all configurations by optimizing their coefficients

Correlation energy ≡ energy reduction
Configuration interaction (CI) is formally exact, but limited to 10 - 20-atom problems

- Hence density functional theory (DFT) both exchange & correlation accounted for.

Hohenberg and Kohn (1964) $V_{BO}(\{X_I\}) = V_{BO}$

$$\updownarrow \qquad\qquad\qquad \nearrow\!\!\swarrow \rho$$

$$v(\underline{x}) = \sum_I \frac{e^2 Z_I}{x_i - X_I} \longleftrightarrow \Psi_G(x_1, \dots, x_n)$$

two non-trivially different ion-electron potentials cannot give rise to the same density and vice versa ⟹ one-to-one marking } 8×10^{60} bytes ⟷ $8 * 10^3$ bytes
compression

Kohn and Sham (1965): fictitious noninteracting electron system

$$V_{BO}(\rho(x)) = V_{independent}(\rho(x)) + V_{Hartree}(\rho(x)) + V_{\substack{exchange\\correlation}}(\rho(x))$$

this is not a "single-determinant" method like H-F.

→ approximate* $V_{exch-corr}$ by a density with Monte Carlo quantum method. derived for homogeneous electron gas.

* LDA: local density approximation
GGA: generalized gradient approximation. } limitations = heterogeneity

hybrid functionals mix DFT and Hartree-Fock (but not really ab initio ?!)

- Semi-empirical electronic-structure methods

use ab initio or experimental information to fit intrinsic electronic quantities

- interatomic potential / force field method: parameterized V_{BO} surface
$O(N)$ linear scaling algorithms.